FINAL TECHNICAL REPORT

GPS Rapid Post-Processing for Earthquake Response

Award Number:	**G14AP00040**
Start Date & End Date:	4/2014 – 4/2015
Principal Investigator:	Ingrid Johanson
Email Address:	ingrid@seismo.berkeley.edu
Institution and Address:	University of California, Berkeley
	Berkeley Seismological Laboratory
	215 McCone Hall # 4760
	Berkeley, CA 94720-4760

Abstract

Rapid access to reliable information is critical for emergency response. In the case of a major earthquake, the mobilization of local, state, and federal disaster operations can be greatly enhanced by dependable, near real-time estimates of location, magnitude, mechanism, and extent of strong-ground shaking. This information can be used to identify endangered communities, to evaluate the impact on lifelines, and to provide input for damage and loss estimation. Current applications of rapid earthquake information include the emergency services, transportation, utilities, telecommunications, and insurance industries. Over the last several years, "ShakeMaps" - representations of the distribution of earthquake peak ground motion and intensity - have emerged as critical products for emergency response operations. However, it is clear that a simple parameterization of hypocenter and magnitude alone does not provide sufficient information for post-earthquake response. This is particularly true for large events, where the finite extent of the source becomes important.

As a result of previous NEHRP funding, we have developed a strategy to perform rapid post-processing (RPP) of GPS data to obtain static offsets and finite fault parameters within 15-20 minutes following an earthquake. RPP provides more accurate offset measurements that real-time processing and can be used to measure smaller earthquakes that would be possible with real-time processing, allowing us to use GPS data to provide better information to earthquake response products on the 15-20 minute timescale. Under this project we sought to make improvements to the nonlinear search process for the fault parameters. These had been performed using a nonlinear search algorithm in Matlab and are highly susceptible to finding local minima. This has now been replaced with a Markov-Chain Monte Carlo sampling method that does not require a starting model and also provides adequate sampling to populate a probability density function for the model parameters, thereby providing an estimate of how well each is determined.

GPS Rapid Post-Processing for Earthquake Response

Award Number: **G14AP00040**

Ingrid Johanson
University of California, Berkeley
Department of Earth and Planetary Science
307 McCone Hall
Berkeley, CA 94720-4767
Telephone: (510) 545-3525;
Fax:(510) 643-5811;
E-mail: ingrid@seismo.berkeley.edu

Introduction

Rapid access to reliable earthquake information is critical for emergency response. In the case of a major earthquake, the mobilization of local, state, and federal disaster operations can be greatly enhanced by dependable, near real-time estimates of location, magnitude, mechanism, and extent of strong ground shaking. Over the last decade, ShakeMap - a representation of the distribution of earthquake peak ground motion and intensity - has emerged as a critical product for emergency response operations, and is valued by emergency response personnel [*Wald et al.*, 1999; 2008]. It provides information that can be used to identify endangered communities, to evaluate the impact on lifelines, and to provide input for damage and loss estimation. Current applications of rapid earthquake information include the emergency services, transportation, utilities, telecommunications, and insurance industries.

ShakeMaps, originally developed as part of the TriNet Project in southern California [*Wald et al.*, 1999], combine observed ground motions (PGA, PGV, and spectral acceleration) with attenuation relations to produce maps showing the distribution of shaking. The attenuation-based shaking model is used to fill in where there is less data and to interpolate between data points and is by default calculated from the distance to the earthquake hypocenter. However, for large earthquakes a simple parameterization of shaking based on hypocentral distance does not provide complete information for post-earthquake response. The finite extent of the source and the effects of directivity become important for larger events. Accounting for finite fault effects, through line source and finite fault modeling (Figure 1), has provided a marked improvement in ShakeMap in many cases [*Dreger et al.*, 2005]. ShakeMap can also use the finite fault information directly. Given rupture dimensions, ShakeMap can instead calculate station distance to the fault plane, which allows higher levels of shaking throughout the rupture area. Including a finite fault plane in ShakeMap already happens for large earthquakes, however it is not yet automated and requires manual intervention to determine the rupture plane and add it to the pro-

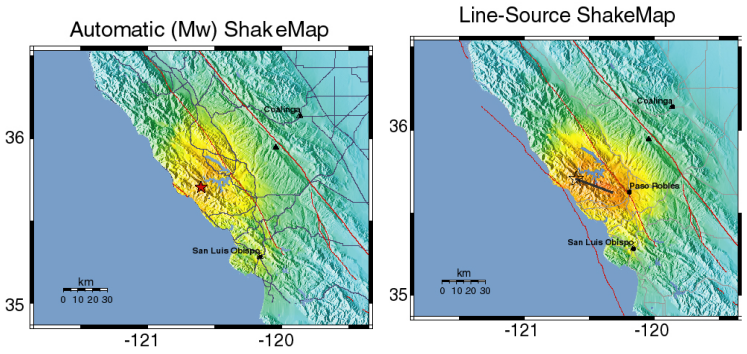

Figure 1. Comparison of ShakeMaps for the 2003 M_W6.5 San Simeon earthquake. Shown are rapid determinations based on only a few seismic stations. Left, standard point-source attenuation models; right, models that include adjustments based on the automated line-source model. The finite source model extends and increases the predicted intensities to the southeast, in agreement with observations.

cessing. With real-time GPS data and rapid post-processing, finite fault parameters are automatically determined and can be provided to ShakeMap within 20 minutes.

Seismic monitoring in California and rapid earthquake information

In Northern California, the University of California, Berkeley Seismological Laboratory (BSL) and the USGS Menlo Park collaborate closely to provide timely and reliable earthquake information to the federal, state, and local governments, to public and private agencies, to researchers nation- and worldwide and to the general public. This collaboration forms the Northern California Earthquake Management Center (NCEMC) of the CISN. The CISN is a partnership among the USGS, California Geological Survey, UC Berkeley, and Caltech, combining efforts in northern and southern California as well as weak and strong motion networks. The NCEMC uses the ANSS Quake Management System (AQMS, formerly called the CISN software) to routinely produce estimates of earthquake location and magnitude within 2-4 minutes after an event. For events of magnitude 3.5 and larger, the CISN centers also produce ShakeMaps within 6-8 minutes.

In 2010 the BSL signed a MOU with the USGS, Menlo Park to share GPS data and to cooperate in providing information to earthquake response products from GPS. An important aspect of NCEDC is that the BSL and USGS processing centers mirror each other, providing mutual back-ups in the case that either center goes offline. In order to incorporate GPS information into CISN, similar mirrored processing will need to be in place for GPS data.

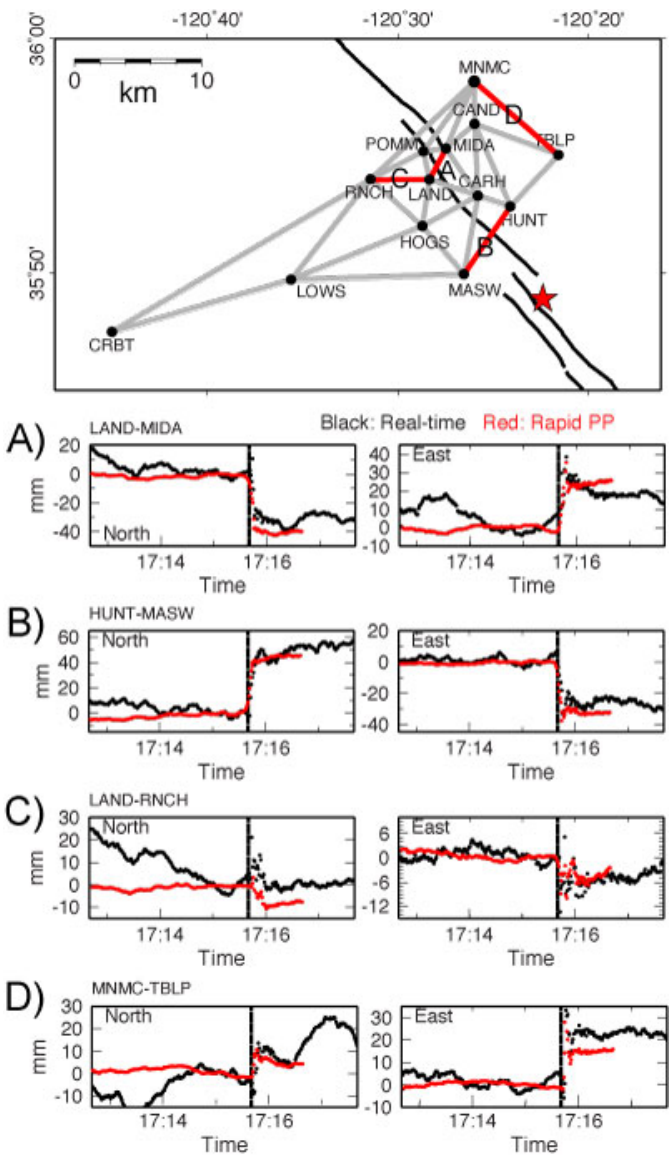

Figure 2: A comparison of simulated real-time (black) and rapid post-processing (red) GPS time series for four baselines during the 2004 Parkfield earthquake. Top panel shows the triangulated processing network in grey, with the four baselines shown highlighted in red. The vertical line shows the origin time of the earthquake.

Geodetic Observations

Although seismic networks can be used to estimate finite-fault parameters, they often have poor sensitivity to the geometry of the rupture. Geodetic networks, particularly if they have stations close to the event, often have much better sensitivity to the orientation and areal extent of rupture, providing a complementary data source for independently estimating rupture parameters of M>6 events. High-rate (1 Hz sampling) GPS data also provide information at low frequencies and for static offsets, information that is difficult to obtain from seismometers [*Larson et al.*, 2003].

Geodetic measurements of coseismic displacements provide important constraints on earthquake faulting, including the location and extent of the rupture plane, unambiguous resolution of the nodal plane, and the distribution of slip on the fault unbiased by rupture velocity assumptions [e.g. *Murray-Moraleda and Simpson*, 2009; *Johanson and Bürgmann*, 2010]. Seismic and geodetic observations can also be inverted simultaneously to improve resolution of finite-fault earthquake source models [*Rolandone et al.*, 2006; *Kim and Dreger*, 2008]. These and other studies typically find that the geodetic data is most sensitive to the fault geometry and provides a smooth slip distribution, and that the seismic data resolves the temporal evolution of the slip including rupture and slip velocity variations, and smaller scale heterogeneity in the slip distribution [e.g. *Cohee and Beroza*, 1994]. For example, *Chi et al.* [2001] found that GPS data were critical for constraining the complex geometry of the northern end of the rupture zone in the Chi-Chi earthquake where the rupture bends to the east.

Rapid Post-processing

Under a previous NEHRP award (G12AP20097), we developed a methodology to use GPS data to provide static offsets and finite fault parameters that can be used for ShakeMap. While

we originally planned to use real-time processing techniques, we soon realized that more precise results could be obtained with high-rate post-processing. Even by waiting as little as 1 minute, better estimates of static offset and smaller offsets can be obtained than with real-time processing. Rapid post-processing (RPP) requires waiting 1-2 minutes after the earthquake for data to accumulate, but displacement time series can then be generated within 5 minutes using Track, a freely-available, open-source processing package developed at MIT. From these, full fault plane determination can be performed within another 5 minutes. While real-time processing techniques are critical for using GPS data for Earthquake Early Warning, the more accurate RPP allows GPS data to be used for smaller earthquakes and still finishes within a time frame appropriate for ShakeMap.

For all processing, we adopt a triangulated network strategy (e.g. grey lines in upper frame of Figure 2), where each baseline pair is individually processed using Track to determine the relative offset between two stations. This strategy makes the processing highly parallel, such that there is a minimal increase in time associated with individual pair processing. It also makes our network resilient against a data outage at any individual station and against the failure of any single processing instance. We have also tailored our fault parameter search to use these relative offsets directly, without performing a network adjustment that would cast each station's offset as relative to a single reference. Altering the model set-up to accept multiple relative offsets is fairly simply accomplished for a linear inverse problem. The major change is to cast the problem $\mathbf{Gm=d}$ as $(\mathbf{G_1-G_2})\mathbf{m}=(\mathbf{d_1-d_2})$.

Using Baysian statistics for geometry determination

Previously the best-fitting geometry was found using a nonlinear search algorithm using functions in Matlab. This process was very sensitive to starting geometry, such that in order to get a robust outcome, several starting geometries needed to be tested and then compared in order to find the true "best-fitting" geometry. This points to uncertainty in the determined model parameters that belies to determination of a single best-fitting geometry. A Bayesian approach to evaluating the set of possible fault geometries provides a quantitative determination of *a posteriori* model parameter probability distributions. While a single final geometry is still sought, the parameter probability distributions allow a quantitative determination of how well determined that geometry is.

We use the MCMC Hammer toolbox for Matlab, which combines a Markov-Chain Monte Carlo (MCMC) sampling process with Bayesian statistics [*Anderson and Segall*, 2013]. A two-stage approach is used for rapid automated processing. In the first stage, the MCMC sampling is harnessed with large step sizes to provide a coarse, but hopefully complete sampling of the probability space in order to quickly determine the overall minimum in data misfit. The initial parameter space is intentionally left very large, with any possible strike or dip being allowed. The limits to the length and width of the fault plane are initially set to twice the average length and width of a fault for the given initial magnitude, as determined by *Wells and Coppersmith* [1994]. The best-fitting fault geometry from the first stage is used as a starting model for the second stage. Also, the search limits in the second stage are restricted to the 95% confidence intervals of the parameter distributions from the first stage. The goal of the second stage is to provide refined sampling of the model parameter space around the best-fitting geometry. The probability distribution from the refined stage will then be used to determine whether the geometry should be automatically passed to downstream applications or not.

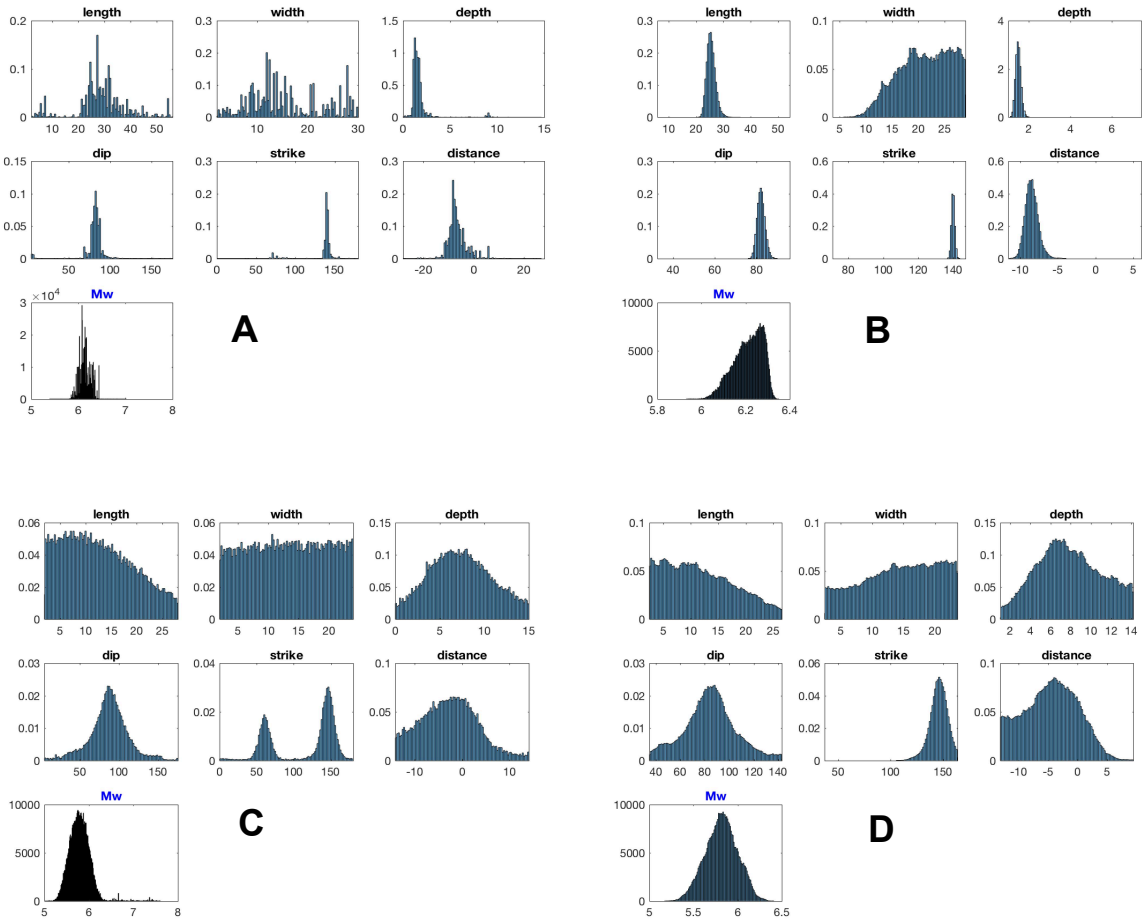

Figure 3: Model Parameter probability density functions. A) After stage 1 (coarse) for the Parkfield earthquake. B) After stage 2 (refined) for the Parkfield earthquake. C) After stage 1 (coarse) for the Alum Rock earthquake. D) After stage 2 (refined) for the Alum Rock earthquake. All distance units are kilometers. “Distance” refers to along strike distance from the hypocenter.

Overall, the results of using the MCMC algorithm are very similar to the nonlinear search algorithm. However, this method provides a much more complete picture of the uncertainties in the final model parameters.

The final fault plane for the Parkfield earthquake is very similar to that determined using the nonlinear search algorithm. The probability density functions further indicate that the majority of the model parameters are well constrained by the data. The exception to this is the width of fault plane (Figure 3b). Geodetic data can struggle to constrain the depths of fault structures, so it is no surprise that there is essentially an even probability of the fault plane being 20 km or wider. This results illustrates the importance of calculating these probability distributions, as any downstream application that is sensitive to fault width would want to be cautious about applying the values determined here.

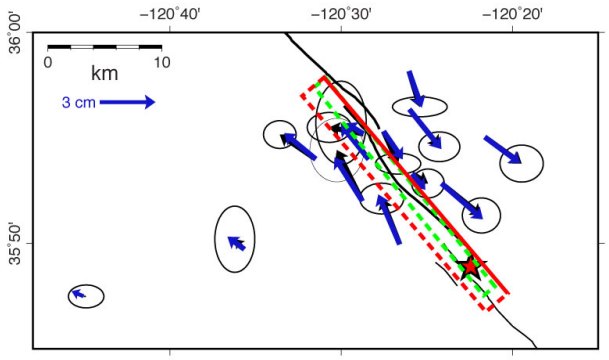

Figure 4: Comparison of fault geometry determined by multiple parallel nonlinear searches (green dashed line) with that determined using MCMC algorithm (red dashed line). Solid red line indicates the shallow edge of the geometry. Data (black vectors) and modeled fits (blue) are shown adjusted assuming a stable network centroid and are shown with 1-sigma error ellipses.

The probability density functions for the Alum Rock earthquake similarly confirm what was qualitatively apparent from the nonlinear geometry search. Namely, that the model parameter for the Alum Rock earthquake are much less well constrained than those for the Parkfield earthquake. This is due to smaller amount of data available and the fact that this data is over a wider area than that available for the Parkfield earthquake. Notable from this analysis is the double peaks in the probability density function for the strike of the fault plane (Figure 4a). The MCMC algorithm finds well-fitting planes both parallel to the strike of the Calaveras fault in the area and conjugate to it. This is similar to what was inferred from the nonlinear geometry search, which was that this event was not fundamentally distinguishable from a point source. Following the MCMC analysis it is clear that many of the geometry parameters are largely unconstrained. This would valuable input for anyone wishing to apply this geometry rapidly after an earthquake. Such distributions would be an indication not to use this geometry further.

Results for the 2014 South Napa Earthquake

On August 24, 2014 a Mw6.0 earthquake occurred on the West Napa fault. This earthquake was one of the first tests of the ShakeAlert earthquake early warning system on a moderate sized earthquake. This included the geodetic early warning algorithm G-larms, which was operational at the time [*Grapenthin et al.*, 2014]. The continuous GPS stations in the area are a mix of those

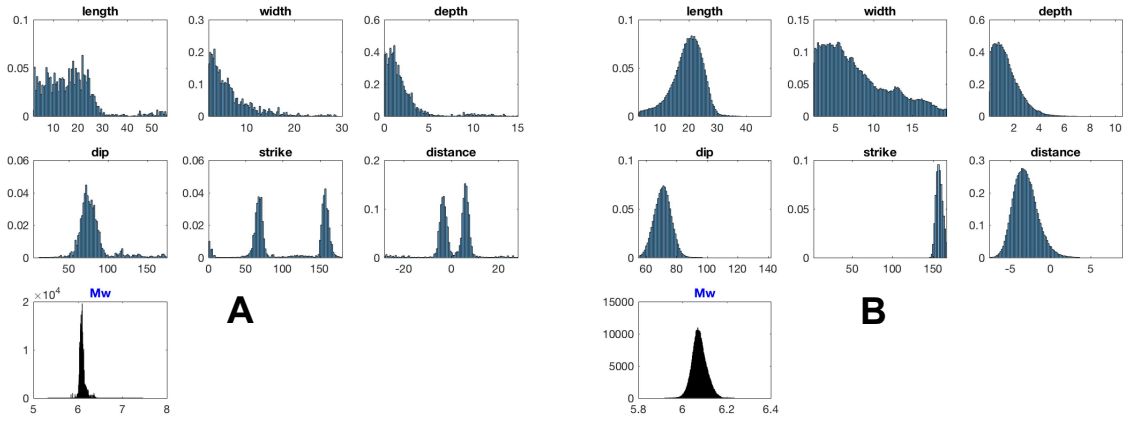

Figure 5: Model Parameter probability density functions for the South Napa earthquake. A) After stage 1 (coarse). B) After stage 2 (refined). All distance units are kilometers. "Distance" refers to along strike distance from the hypocenter.

operated by UNAVCO as a part of the Plate Boundary Observatory (PBO) and stations from the Bay Area Regional Deformation (BARD) network operated by the BSL. Unfortunatly several PBO stations were not available via real-time stream at that time, including the closest station to the epicenter, P261. Likewise P199, P200 and P264 also did not have real-time data streams available (Figure 4). G-larms considers a limited set of fault geometries before determining the size and extent of slip from the real-time streams. This helped compensate for the lack of near-field data and allowed it to produce a magnitude estimate of Mw 5.9; a value quite close to the final preferred value of Mw 6.0.

The goal of the methodology presented here is to determine the fault plane geometry. This was not feasible given the unavailable near-field data. The MCMC algorithm shows that nearly all the model parameters were largely unconstrained for the scenario where only real-time stations were used. If the closer stations had been available in real-time, the outcome would have been much better and this is the scenario presented here.

The model parameter probability density functions for the coarse MCMC sampling (stage one) show two families of solutions: the San Andreas fault parallel and conjugate fault planes. Like the results for the Alum Rock earthquake, the San Andreas fault parallel family has a smaller data misfit and so is preferred and passed along to the refined MCMC sampling (stage two). The refined probability density functions show nicely constrained distributions for most of the values, except for the fault plane width. Similar to the results for the Parkfield earthquake, the width of the fault plane is the least well-constrained parameter. While a shorter slip patch is preferred, deeper slip cannot be ruled out.

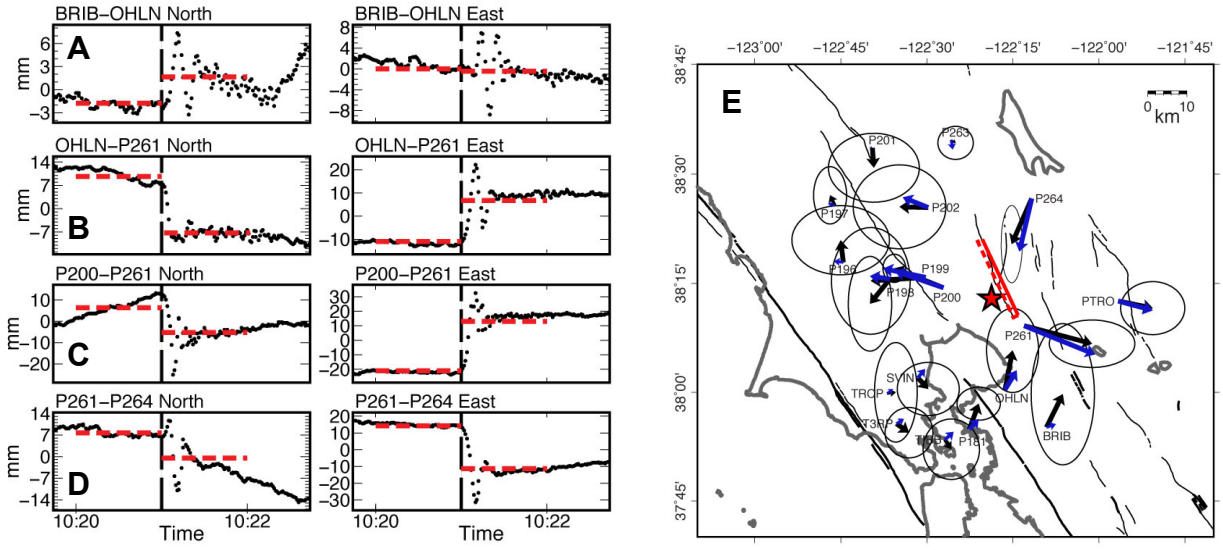

Figure 6: A-D) Time series (black) and calculated offset (red) for individual station pairs from the South Napa earthquake. E) Data (black) and modeled (blue) station offsets adjusted assuming a stable network centroid; shown with 1-sigma error ellipses. Red dashed box is the outline of the determined fault plane, the solid line is the shallow edge of the slip patch. Red star marks the earthquake epicenter.

Conclusions

Use of the MCMC algorithm to determine the fault plane geometry on a "rapid" timescale is a significant improvement over the multiple, parallel nonlinear searches previously used. The method provides robust, repeatable results, that agree well with both post-event models and results using the nonlinear search algorithm. Furthermore is does this with little additional time required. The MCMC runs typically take 5-10 minutes to run, well within the time constraint of rapid post-processing. The additional information provided by the probability density functions is valuable for assessing the quality of the solution, which depends on having an earthquake large enough to generate measurable offsets and stations close enough to be sensitive to the finite nature of the rupture.

1-sigma (68% confidence) ranges	Parkfield EQ	Alum Rock EQ	South Napa EQ
Length (km)	24.0 – 27.1	4.7 – 19.2	14.4 – 24.6
Width (km)	15.7 – 26.7	6.7 – 20.6	3.6 – 12.9
Top Depth (km)	1.3 – 1.6	3.7 – 10.7	0.4 – 2.4
Dip (degrees)	80 – 84	62 – 108	65 – 76
Strike (degrees)	139 – 140	137 – 154	153 – 161
Distance from Hypocenter (km)	-9.3 – -7.6	-9.4 – 0.9	-4.5 – -1.6
Mw	6.1 – 6.3	5.6 – 6.0	6.0 – 6.1

Table 1: Sample 1-sigma ranges for parameters determined for various earthquakes. Values of dip greater than 90 degrees indicate a strike greater than 180 degrees.

References

Anderson, K., and P. Segall (2013), Bayesian inversion of data from effusive volcanic eruptions using physics-based models: Application to Mount St. Helens 2004-2008, *J. Geophys. Res.*, *118*(5), 2017–2037, doi:10.1002/jgrb.50169.

Chi, W. C., D. Dreger, and A. Kaverina (2001), Finite-Source Modeling of the 1999 Taiwan (Chi-Chi) Earthquake Derived from a Dense Strong-Motion Network, *Bull. Seismol. Soc. Am.*, *91*(5), 1144–1157, doi:10.1785/0120000732.

Cohee, B. P., and G. C. Beroza (1994), Earth-prints: A comparison of two methods for earthquake source inversion using strong motion seismograms, *Annals of Geophysics*, *37*(6), 1515–1538.

Dreger, D. S., L. Gee, P. Lombard, M. H. Murray, and B. Romanowicz (2005), Rapid Finite-source Analysis and Near-fault Strong Ground Motions: Application to the 2003 Mw 6.5 San Simeon and 2004 Mw 6.0 Parkfield Earthquakes, *Seismological Research Letters*, *76*(1), 40–48, doi:10.1785/gssrl.76.1.40.

Grapenthin, R., I. A. Johanson, and R. M. Allen (2014), The 2014 Mw 6.0 Napa earthquake, California: Observations from real-time GPS-enhanced earthquake early warning, *Geophysical Research Letters*, *41*(23), 8269–8276, doi:10.1002/2014GL061923.

Johanson, I. A., and R. Bürgmann (2010), Coseismic and postseismic slip from the 2003 San Simeon earthquake and their effects on backthrust slip and the 2004 Parkfield earthquake, *Journal of Geophysical Research*, *115*(B7), B07411, doi:10.1029/2009JB006599.

Kim, A., and D. S. Dreger (2008), Rupture process of the 2004 Parkfield earthquake from near-fault seismic waveform and geodetic records, *Journal of Geophysical Research*, *113*(B7), doi:10.1029/2007JB005115.

Larson, K., P. Bodin, and J. Gomberg (2003), Using 1-Hz GPS data to measure deformations caused by the Denali fault earthquake, *Science*, *300*(5624), 1421–1424.

Murray-Moraleda, J. R., and R. W. Simpson (2009), Geodetically Inferred Coseismic and Postseismic Slip due to the M 5.4 31 October 2007 Alum Rock Earthquake, *Bull. Seismol. Soc. Am.*, *99*(5), 2784–2800, doi:10.1785/0120090017.

Rolandone, F., D. Dreger, M. Murray, and R. Bürgmann (2006), Coseismic slip distribution of the 2003 Mw6.6 San Simeon earthquake, California, determined from GPS measurements and seismic waveform data, *Geophysical Research Letters*, *33*(16), L16315, doi:10.1029/2006GL027079.

Wald, D. J., V. Quitoriano, T. H. Heaton, H. Kanamori, C. W. Scrivner, and C. B. Worden (1999), TriNet "ShakeMaps": Rapid generation of peak ground motion and intensity maps for earthquakes in southern California, *Earthquake Spectra*, *15*(3), 537–555.

Wald, D., K.-W. Lin, K. Porter, and L. Turner (2008), ShakeCast: Automating and Improving the Use of ShakeMap for Post-Earthquake Decision-Making and Response, *Earthquake Spectra*, *24*(2), 533–553, doi:10.1193/1.2923924.

Wells, D. L., and K. J. Coppersmith (1994), New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement, *Bull. Seismol. Soc. Am.*, *84*(4), 974–1002.